

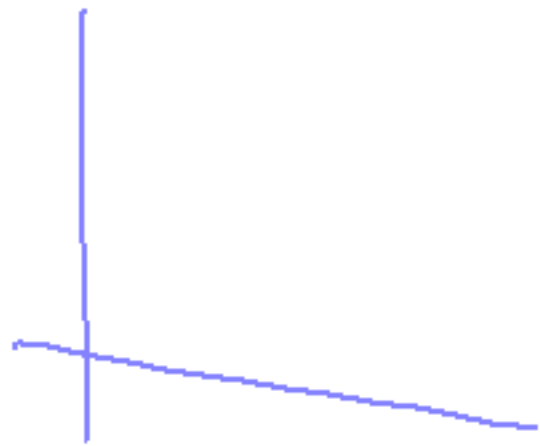
Ch. 10 Review

① Explanatory - x -variable
* usually researcher sets

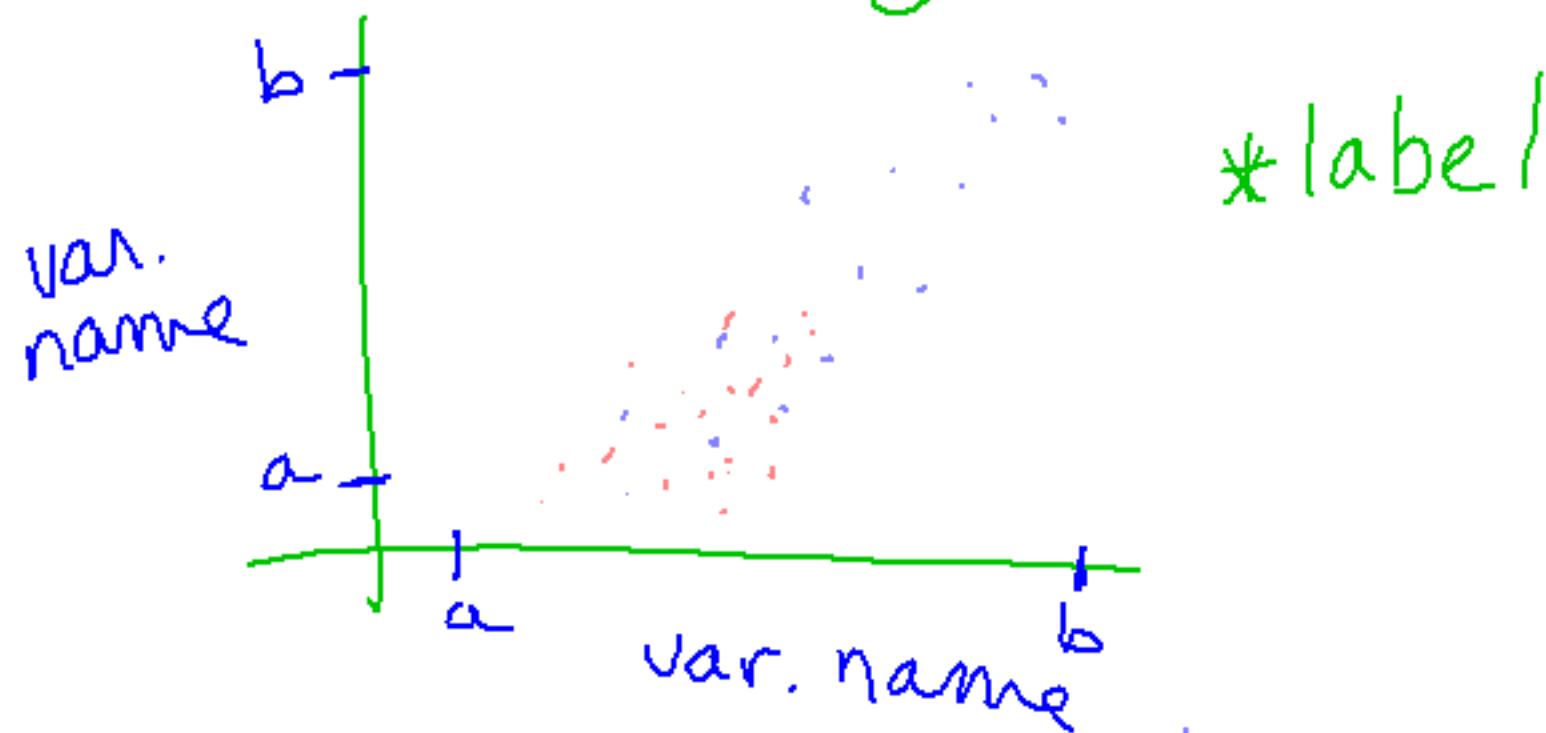
Response - y -variable

* usually measured by
researcher

* change because of explan. var.



② Scatterplot - 2 quantitative
* categorical: by diff. colors/shapes



③ Describe: Form: linear/curved
Direction: + or -
Strength: weak, mod, strong

④ LSR line:

Least Squares
Regression
line

$$\hat{y} = a + bX$$

intercept slope

⑤

$$\hat{y} = b_0 + b_1 X$$

* formulas on formula sheet

$$a = \bar{y} - b\bar{X}$$
$$b_0 =$$

$$b = r \frac{S_y}{S_x}$$
$$b_1 =$$

⑥ r = correlation coefficient

* how linear (+ direction) the relationship btw. x and y is.

$$-1 \leq r \leq 1$$

close $-1, 1$ = strong linear

close to 0 = weak linear

$$r = 0.78 \quad r = 0.2$$

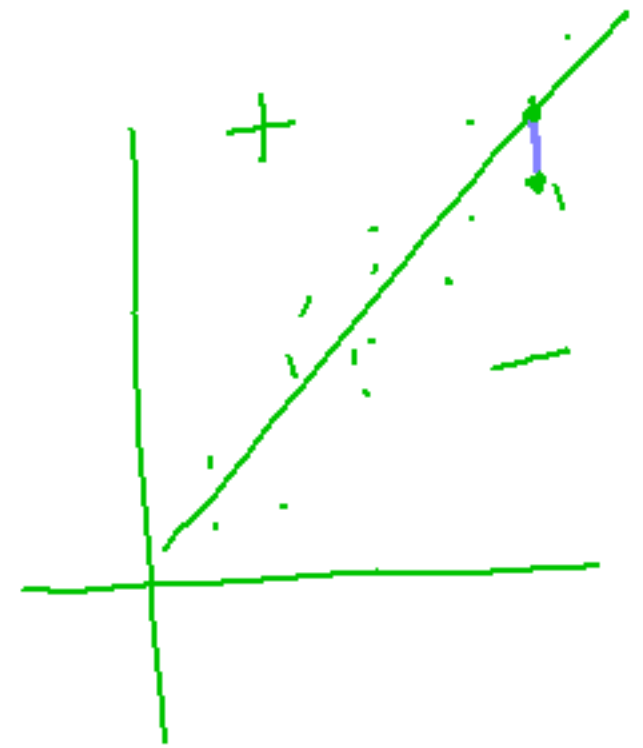
+
linear
mod. strong



* ⑦ r^2 = coefficient of determination

_____ % of the ^{variation} change in y-variable
that is due to the change in x-var.
the LSR line

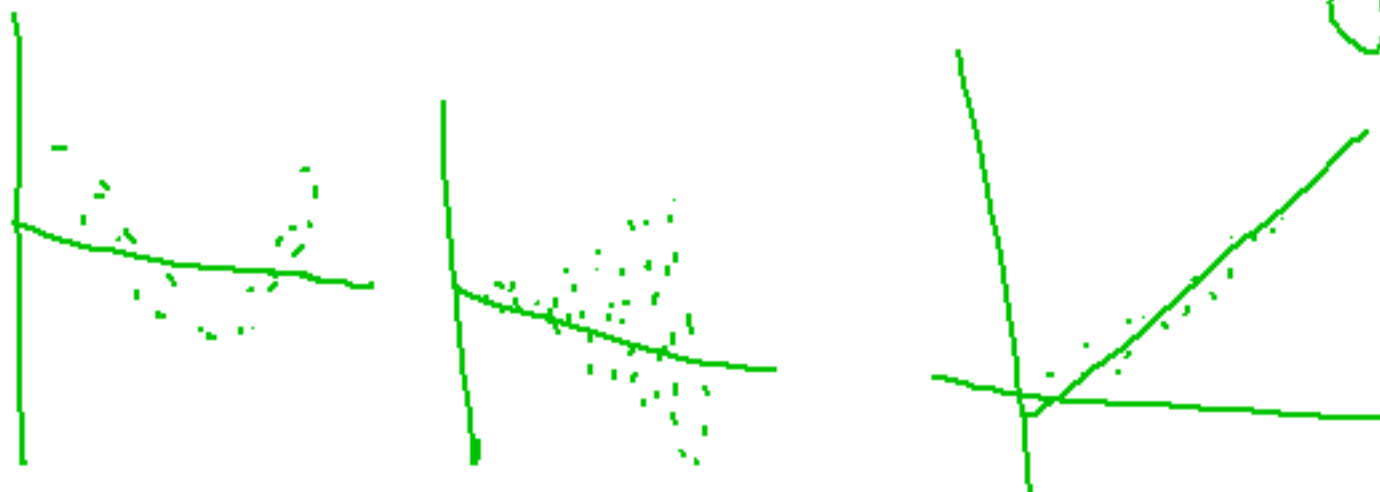
⑧ residual = error in y-direction
= observed - predicted
actual data from LSR line





Scattered = good . line is a good fit for data

Pattern = line is not a good fit for data



Do worksheet on Turnpike Tolls. Add:

#9 Create resid. plot + describe

$$\hat{y} = -0.1157 + 0.04(150)$$

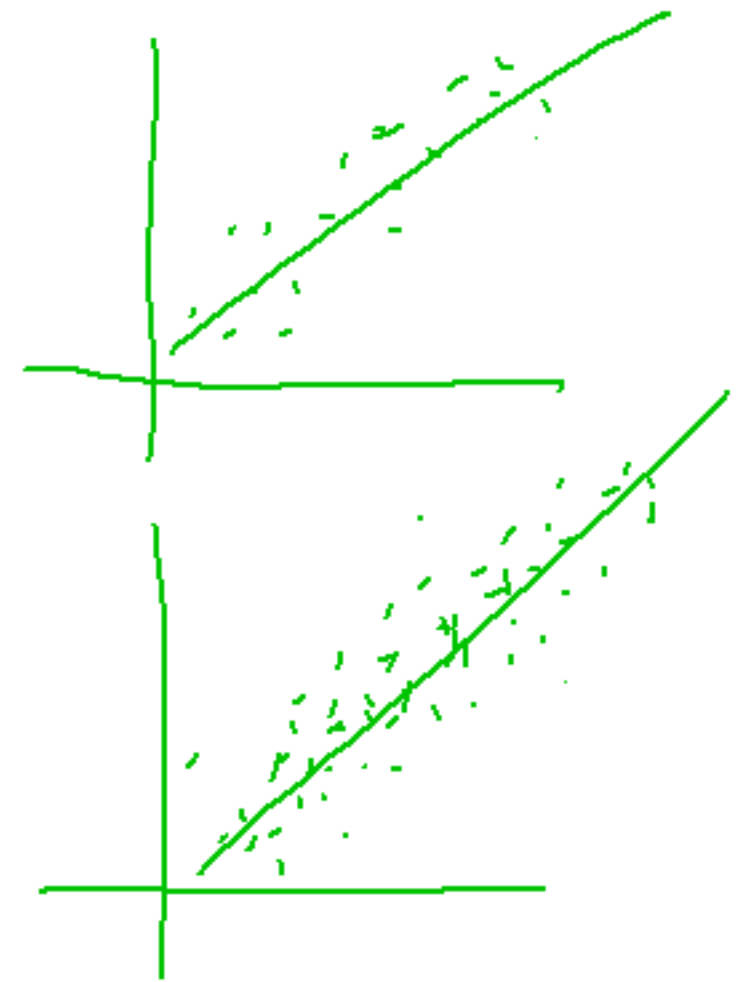
$$\hat{y} = \$5.91$$

More review... LSR line

sample: $\hat{y} = a + bx + e_i$
0.04

pop: $y = \alpha + \beta x + \epsilon_i$

a	α	intercept
b	β	slope
e_i	ϵ_i	resid/errors
\hat{y}	y	resp. var.



Ch. 10: Inference on B

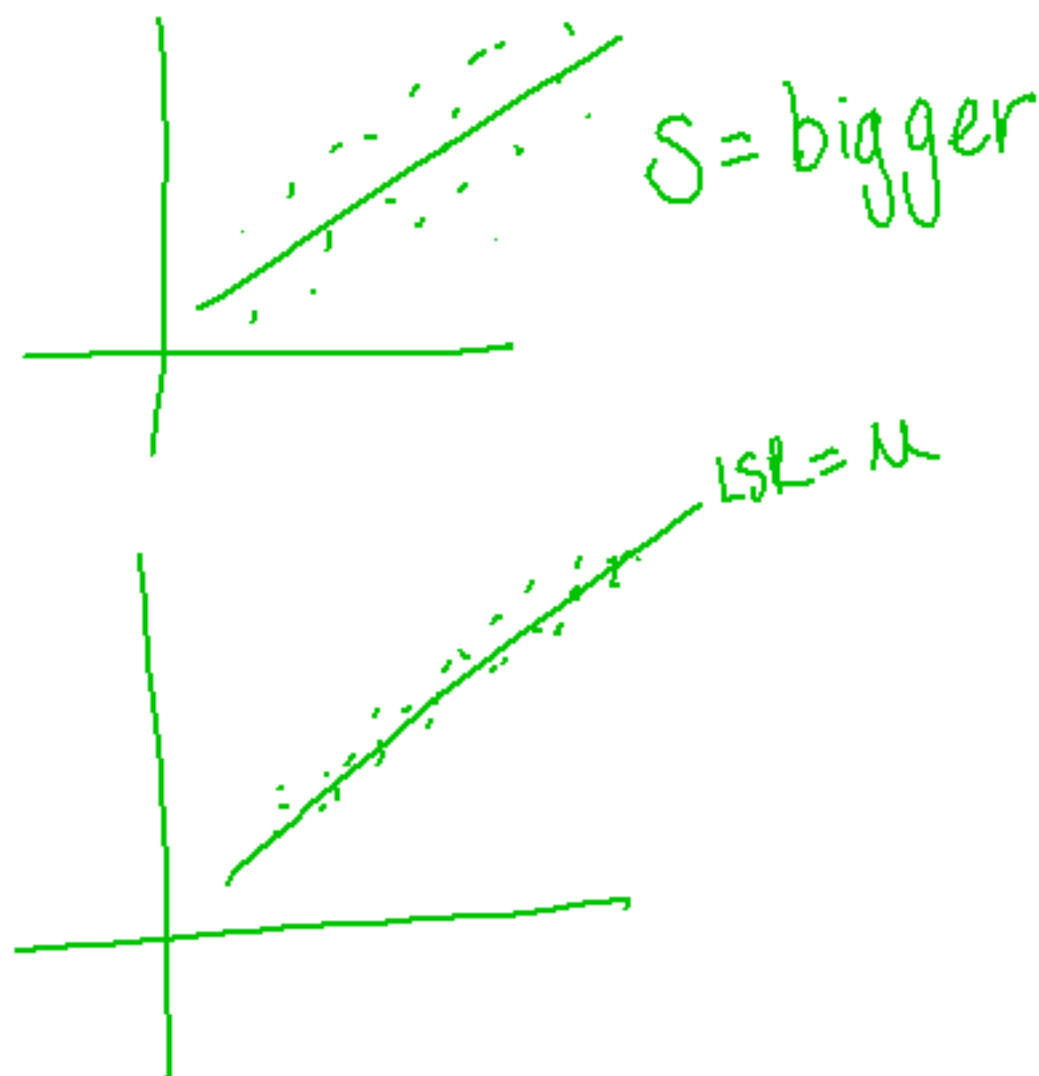
$$\sum \text{resid} = 0$$

$$\sum \text{resid}^2 = 0$$

$$\sigma_{\text{resid}} = ?$$

/
estimate with S

↑
from sample



* testing the slope of population regression line

$$H_0: \beta = 0$$

$$H_a: \beta \neq 0$$

$$t = \frac{b - \beta}{SE_b}$$

$$P(t \geq \text{test stat} \mid df = n - 2)$$

data pts.

Concl: •

• We have suff. evid. that the slope of population regression line is _____ to 0.

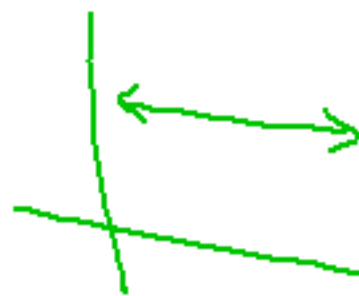
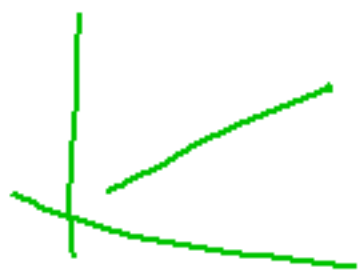
• Thus... as x-variable increases, the y-variable _____.

+ increases

- decrease

≠ changes

③ doesn't change



Conf. Int:

$$b \pm t^* SE_b = (a, b)$$

← from t table

$$df = n - 2$$

conf level on bottom

We are ____ % conf. that the

slope of population regression line

btw. x and y is btw a and b $\frac{y\text{-units}}{x\text{-units}}$.

Assump

① 2 independent SRS

② true relationship
is linear

Check

② assumed

① Computer Output

Intercept
Constant

Coeff

Std. Dev.

teststat

P-val

a

~~SE_a~~

~~X~~

~~X~~

b

SE_b

t

p-value

Year

↗
name of
X-var.

S =

r² =

~~r²_{adj} =~~

↖ std. dev.
of resid.

$$\hat{y} = -5.76 + 0.348x$$

